

A Deep Learning Method for Automatic Visual Attention Detection in Older Drivers

Belkacem Chikhaoui¹, Perrine Ruer², and Évelyne F. Vallières¹

¹ LICEF Research Center, TELUQ University, QC, Canada,
belkacem.chikhaoui@teluq.ca, pruer.pro@gmail.com,
Evelyne.Vallieres@teluq.ca

² HEC Montreal, Montréal, Québec, Canada

Abstract. This paper addresses a new problem of automatic detection of visual attention in older adults based on their driving speed. All state-of-the-art methods try to understand the on-road performance of older adults by means of the Useful Field of View (UFOV) measure. Our method takes advantage of deep learning models such as Long-short Term Memory (LSTM) to automatically extract features from driving speed data for predicting drivers' visual attention. We demonstrate, through extensive experiments on real dataset, that our method is able to predict the driver's visual attention based on driving speed with high accuracy.

Keywords: UFOV, deep learning, LSTM, classification, divided attention, older drivers.

1 Introduction

Visual and cognitive abilities are important parameters for safe driving. These abilities tend to decrease naturally with aging, and many older adults become unable to drive because of a serious decline in their visual and cognitive abilities. The decline in visual and cognitive abilities may lead to unsafe driving [1].

One of these abilities concerned divided attention, which is defined as the ability to construct information from multiple sources that are critical to the execution of a specific task (e.g., driving) [1], has received particular interest from researchers as a good indicator of driving performance in older drivers [1, 2]. In fact, in the context of driving, drivers face multiple stimuli coming from the environment including the car, the road, other drivers, the weather, and driving time of day. These stimuli create a complex context, particularly for older drivers, that requires from them to shift attention between all these stimuli in order to ensure safe driving.

The Useful Field of View (UFOV) [3] is a well known and widely used test to measure visual attention [4]. It consists of three subtests of visual attention: processing speed, divided attention, and selective attention. The first one is to identify peripheral targets, the second and third subtests are to identify peripheral target presented either in the presence or absence of distractors while completing a central discrimination task [5–7].

The UFOV test has been shown to be highly effective in predicting driving outcomes such as 1) predicting crash risk among older adults [8–10], 2) predicting on-road performance [1, 11], 3) predicting driving in the presence of distracters [12].

The overall studies using UFOV focus on driving outcomes and performance based on obtained UFOV test results. In this regard, our work is original and tries to predict the UFOV divided attention results based on older adults performances during driving, more specifically driving speed. To the best of our knowledge, no prior studies have examined the relationship between driving speed and divided attention. Therefore, the present research was initiated as it may have a great potential in developing assistance systems to help older adults drivers improve their safety while driving. In the research presented in this paper, we resort to deep learning models, particularly Long-Short Term Memory model (LSTM), to automatically detect older adult’s divided attention based on his/her driving speed. The rationale of using LSTM models is threefold: 1) it allows to analyze raw speed data in order to automatically extract important features to perform predictions, 2) it allows to learn features across time by using their internal memory, and 3) it allows to process arbitrary sequences of inputs (speed data) unlike other existing feedforward neural network models. The major contributions of this paper can be summarized as follows:

- Proposing a LSTM based method to automatically detect older adults drivers’ visual attention based on driving speed data.
- Conducting extensive experiments through real data to validate the proposed method.
- Demonstrating the performance and superiority of our proposed method compared to the state-of-the-art methods.

The rest of the paper is organized as follows. First, we give an overview of related work in Section 2. Section 3 describes the proposed model in terms of LSTM representation, and divided attention detection. The results of our experiments on a real driving dataset are presented in Section 4. Finally, Section 5 presents our conclusions.

2 Related work

Multiple studies have been conducted examining the relationship between UFOV test performance and driving. These studies have used a variety of outcome measures, including on-road driving performance, crashes, simulated driving performance, and self-reported driving performance that assess components of driving ability that are neither exhaustive nor necessarily unique.

Driving performance can be assessed by cognitive tests, simulation, on-road driving tests, or a combination of these. In the driving activity, people have to respond rapidly to risks with good abilities. Previous research has investigated specific functions and/or deficits for any correlation with safe driving and in order to predict on-road driving performance.

Some physical impairments found frequently among older people may have negative impacts on their driving skills. Among them, a deficit in visual-cognitive functions, and more specifically divided attention, which is a good predictor of driving performance for older drivers [13].

Aging is associated with decreased performance on the UFOV test due to the higher prevalence in visual search, attention, speed of deficit treatment [13], poorer vision and mental status [14]. Poor performance on the UFOV test was associated with increased crash risk. Older drivers with UFOV impairments are twice as likely to be involved in a road accident [5]. The utility of the UFOV measure has been investigated in different studies. However, a link between the number of crashes and UFOV has been examined in the research [5]. Results showed a significant link between visual attention measures and a car crash for the older drivers. Authors highlighted that self-reported accident frequency by older drivers is not an accurate measure of their actual accident frequency. Authors said the use of this dependent measure in earlier studies may have led to erroneous conclusions about vision and driving relationships [15].

A number of studies used a variety of outcome measures. Some of them investigated on-road driving performance, particularly with simulated driving performance and they assessed driving performance and UFOV test performance. Results reported that both decreased with age. Older is the driver, the more errors are reported during the on-road driving assessment with a slower performance during the UFOV test [16]. Willstrand et al. [16] also highlighted an increasing number of driving errors seen and correlated with age-related reduced selective attention. The reduction of selective attention has a negative effect on driving speed, that is driving too fast, which was the most common error in the on-road driving assessment [16]. Baldock et al. [17] emphasized the importance of speed adaptation, especially the “approach speed”, as it can lead to a delay in the needed braking. The third phase of the UFOV test may be an intervention for older drivers to increase their awareness of attentional problems when driving. Reducing speed when appropriate is an important factor to study with older drivers, as it gives them more time for visual search (processing speed and selective attention [16]).

3 Proposed Method

In order to understand how to detect divided attention based on speed data, firstly we describe the UFOV test in detail. Secondly, we will introduce our proposed method based on LSTM model.

3.1 UFOV test

The UFOV is a computer-administred and computer-scored test of functional vision and visual attention [18], which can be predictive of ability to perform many everyday activities, such as driving a vehicle. The UFOV test lasts about 15 minutes and is highly recommended for drivers having an age of 55 years

or older, particularly those who suffer from health problems including cognitive deficits [18].

UFOV consists in three subtests for assessing speed of visual processing under complex task demands [18]:

1. The first subtest (T1) consists in identifying a target presented in a centrally located fixation box presented for varying lengths of time.
2. The second subtest (T2) consists in identifying a target and localizing a simultaneously displayed target on the periphery of the computer screen.
3. The third subtest (T3) is identical to the second subtest, except that the displayed target is embedded in distractors to make the test more difficult. Each subtest has a threshold score as shown in Table 2.

Table 1 presents an example of UFOV test results (or score) in milliseconds (ms) obtained in real experiments. The range is 17 to 500 ms [18].

Table 1. Example of UFOV test results obtained in real experiments.

Participant	UFOV results		
	Subtest (T1)	Subtest (T2)	Subtest (T3)
1	17	17	197
2	17	23	110
3	17	87	177
4	83	44	287
5	17	23	203
6	20	93	380

The UFOV results can be categorized in different categories called classes of divided attention as shown in Table 2

Table 2. UFOV test values categories

Subtest	Classes		
	Class 1: Healthy	Class 2: Moderate	Class 3: Severe
T1	≤ 30	$>30, \text{and}, \leq 60$	> 60
T2	< 100	$\geq 100, \text{and}, < 350$	≥ 350
T3	< 350	$\geq 350, \text{and}, < 500$	≥ 500

By using this definition of classes, results (from Table 1) are categorized in Table 3.

Table 3. Example of UFOV test results classes.

Participant	UFOV results		
	Subtest (T1)	Subtest (T2)	Subtest (T3)
1	1	1	1
2	1	1	1
3	1	1	1
4	3	1	1
5	1	1	1
6	1	1	2

Once the classes are defined, we will be able to label the dataset accordingly. Therefore, with driving speed data, we try to predict the corresponding class using the LSTM model described in the next section.

3.2 LSTM model

This section introduces the LSTM deep learning model we used to predict divided attention.

LSTM models are a type of recurrent neural networks (RNN) for processing, classifying and making predictions based on times-series data. In contrast to other deep learning models such as Convolutional Neural Networks [19, 20], LSTM models are able to extract temporal features from data. This is a very important characteristic, particularly in driving context where driving speed tend to change over time due to different environmental factors. Moreover, LSTM models have the property of selectively remembering patterns for long durations of time, which make them advantageous compared to basic RNN models.

All recurrent neural networks have the form of a chain of repeating modules of neural network, normally a single tanh layer [21]. In LSTM models, the repeating modules have different structures so that instead of having one single tanh layer, they have four layers as shown in Figure 1 (this figure is taken from [21]).

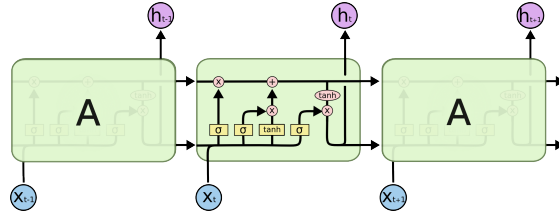


Fig. 1. LSTM structure.

The central component of LSTM model is the cell state presented in Figure 1 with a horizontal line in the top. Information in the cell state could be added or

removed, and this process is regulated by components called gates [21]. In other words, gates have the ability to let or not let information through [21].

To prevent the LSTM model from overfitting, we used the dropout principle [22]. The dropout consists in randomly dropping units with their connections from the neural network during training. Therefore, we randomly dropped 20% of units and their connection after each hidden layer. This percentage is recommended in literature [22]. We experimentally setup the number of units to 256 with a batch size of 128. The number of epochs is experimentally setup to 50. Our LSTM model was implemented in Python programming language using Tensorflow [23].

4 Experimental study

In this section, we describe our experiments conducted at the University of Montreal for data collection and validation of proposed method.

The UFOV test, as mentioned previously, measures three types of attention: selective, shared and divided attention and detects an attention deficit. It is a 15 minutes computer test. The requirements increase in complexity between tasks. The purpose is to evaluate the visual processing speed of a person by detecting, identifying and locating targets briefly presented on the screen. This test is recommended to measure driving ability for older drivers [5, 18].

Twenty-eight participants (14 women, 14 men) aged between 55 and 79 years old (mean: 63,86 y-o; standard-deviation: 6,20) were recruited. They drive on average five days per weeks. All participants have a valid driving licence and have been driving for several years (mean: 45 y-o, sd: 8,95). Participants did not report sleep disturbance or medication use that may affect driving.

All participants went to the driving simulation laboratory at the University of Montreal. Participants received a financial compensation of 60 CAD. The experiments were carried for 20 days during a two-month period.

The research protocol was to provide participants a consent form followed by a first sociodemographic questionnaire. Then, participants performed the UFOV test before driving in the simulator for about 40 minutes. Different data were saved with the simulator as speed, acceleration, braking and steering wheel angle. After driving, they answered "a final post-driving questionnaire" regarding their driving experiment.

The driving simulator is an intermediate level simulator. It is interactive with a fixed base, composed of an entire car (Honda Civic) with fully functional pedals, steering wheel and dashboard as shown in Figure 2.

The simulator is parked in a brightly room in the Department of Psychology at the University of Montreal. The simulator faces a large projection screen (three metres by four metres), which allows a 70-degree view. The road image is computerized and projected by an RGB overhead projector. The location and speed (km/h) of the vehicle on the x, y and z axes are recorded during each simulation. We used these data for our analysis.



Fig. 2. Driving simulator used in our experiments.

5 Classification results

To evaluate our proposed method for automatic detection of the different classes of drivers' divided attention, we divided the data collected into trainset (70 %) and testset (30 %). The trainset is used to build and train the LSTM model whereas the testset is used to validate the model. To evaluate the performance of the LSTM model, we used the precision, recall, and F-measure metrics. The classification results obtained are presented in Table 4.

Table 4. Classification results.

	Classification results		
	Subtest (T1)	Subtest (T2)	Subtest (T3)
Precision	0.9645	0.8581	0.8828
Recall	0.9642	0.8559	0.8733
F-Measure	0.9643	0.857	0.8780

As shown in Table 4, our method shows promising results in classifying the three subtests T1, T2 and T3 with high accuracy. More specifically, the highest performance was observed in T1 with an F-measure equals to 0.9645, while T2 represents the lowest performance with an F-measure equals to 0.857. The rational of having lower F-measure for T2 and T3 compared to T1 is the lack of data about the moderate and severe classes in these two categories. More participants should be recruited in the future in order to overcome this limitation. Confusion matrices obtained for the classification of the three subtests are presented in Figure 3

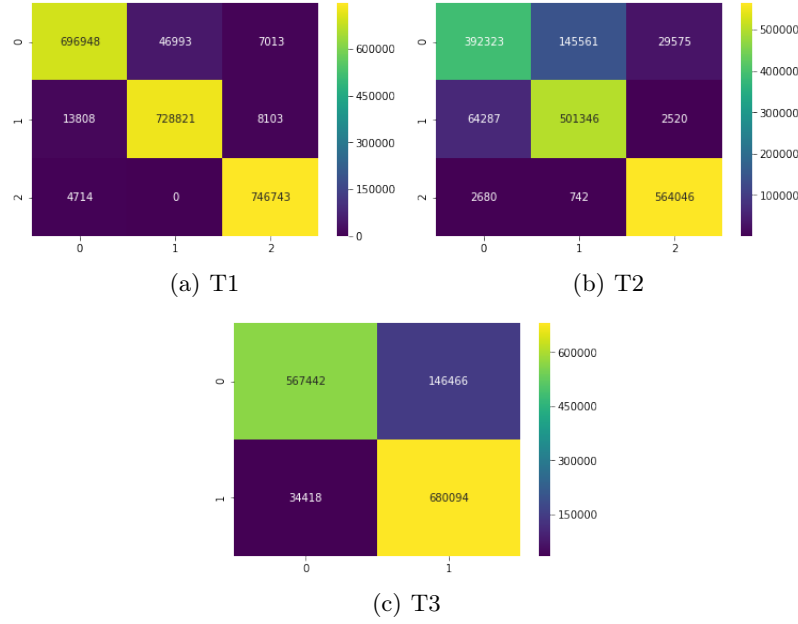


Fig. 3. Confusion matrices.

As shown in Figure 3(a), some samples of class T3 were misclassified as class T1. This could be explained by the similarity in terms of some driving patterns at the beginning of the driving test as shown in Figure 4 compared to T2.

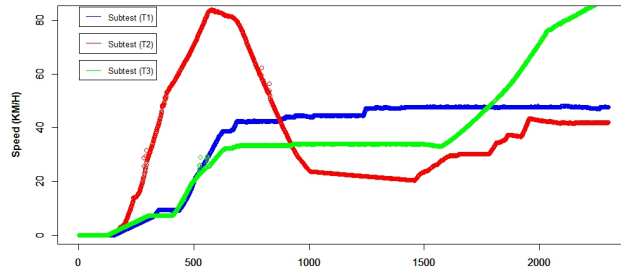


Fig. 4. Speed examples for participants in each class.

We also evaluated the performance of our proposed method compared to the well known state-of-the-art methods of the domain such as decision trees,

naive Bayes, Bayesian networks, support vector machines, K-nearest neighbor, and perceptron neural network. The comparison results for the three different subtests are presented in Figures 5(a), 5(b), and 5(c) respectively.

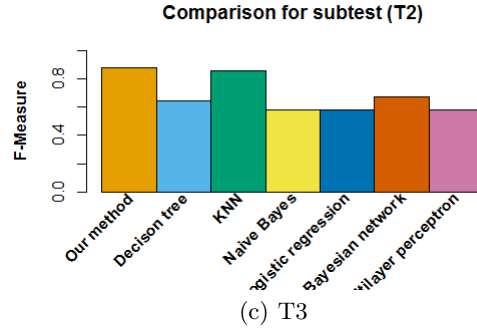
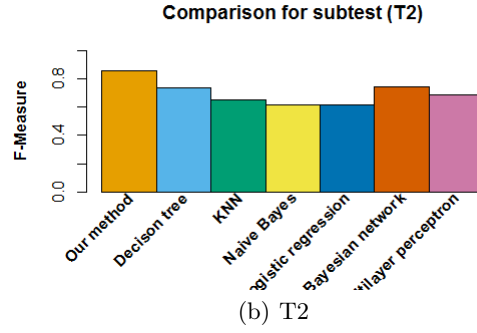
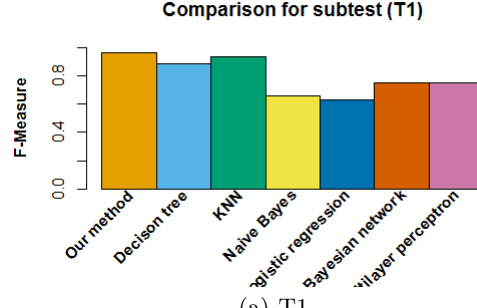


Fig. 5. Classification accuracy using different values of window size.

As shown in Figure 5, our proposed method outperforms all the existing state-of-the-art methods in terms of classification accuracy followed by the K-NN method and Bayesian networks.

Some limits appear in this study. The first one concerns the number of participants. The sample was convenient with only 28 participants coming from the metropole area of Montreal. A further study has to integrate more participants coming from suburban area or countryside to confirm these results. Another limit is the age group. We proposed a solution with only one age group (55 to 79 years old). As people is driving older and older, a comparison should be interesting between 'young' older adults and oldest one (55-74 years-old vs 75-100 years-old).

6 Conclusion

In this paper, we demonstrated the possibility of automatically detecting visual attention in older drivers based on driving speed data and their score from the UFOV test. Our method uses a deep learning model based on LSTM to automatically extract features from raw data and to detect with high accuracy the three different subtests of the UFOV test to determine visual attention of older drivers. We empirically demonstrated the suitability of the proposed method and its superiority compared to the state-of-the-art methods. The results of this research are of great importance for vehicle manufacturers who could integrate our finding to improve safety and security of older drivers. Moreover, the outcome of this research could have a great potential in the future to replace UFOV measure which is not easily reachable for everyone and requires human resources and temporal planning because of its time consuming.

7 Acknowledgement

The authors would like to thank the professors Jacques Bergeron, University of Montreal, and Charles Gouin Vallerand, TELUQ University, to share the data used in this study.

References

1. Crizzle Alexander M Winter Sandra M Lanford Desiree N Sherrilene Classen, Wang Yanning. Predicting older driver on-road performance by means of the useful field of view and trail making test part b. *The American journal of occupational therapy*, 67(5):574–582, 2013.
2. Roenker DL Miller RL Griggs DS Ball KK, Beard BL. Age and visual search: Expanding the useful field of view. *Journal of the Optical Society of America A: Optics and Image Science*, 5:2210–2219, 1988.
3. Owsley C Wood JM. Useful field of view test. *Gerontology*, 60(4):315–8, 2014.
4. Sloane ME. Roenker DL. Bruni JR. Ball K., Owsley C. Visual attention problems as a predictor of vehicle crashes in older drivers. *Investigative ophthalmology and visual science*, 34(11):3110–3123, 1993.

5. Cynthia Owsley, Karlene Ball, Gerald McGwin, Jr, Michael E. Sloane, Daniel L. Roenker, Milton F. White, and E. Todd Overley. Visual processing impairment and risk of motor vehicle crash among older adults. *JAMA*, 279(14):1083–1088, 1998.
6. Cynthia Owsley, Gerald McGwin Jr, and Karlene Ball. Vision impairment, eye disease, and injurious motor vehicle crashes in the elderly. *Ophthalmic Epidemiology*, 5(2):101–113, 1998.
7. Allman RM Ball K Owsley C Sims RV, McGwin G Jr. Exploratory study of incident vehicle crashes among olderdrivers. *The Journals of Gerontology: Series A*, 55(1):22–27, 2000.
8. Kathryn T. Goode, Karlene K. Ball, Michael Sloane, Daniel L. Roenker, David L. Roth, Renee S. Myers, and Cynthia Owsley. Useful field of view and other neurocognitive indicators of crash risk in older adults. *Journal of Clinical Psychology in Medical Settings*, 5(4):425–440, Dec 1998.
9. Vance DE Stavrinou D McManus B, Cox MK. Predicting motor vehicle collisions in a driving simulator in young adults using the useful field of view assessment. *Traffic Inj Prev*, 16(8):818–823, 2015.
10. Dubinsky Richard McDowd Joan M, Atchley Paul. The role of visual attention in predicting driving impairment in older adults. *Psychology and Aging*, 20(4):610–622, 2005.
11. Edwards JD Roth DL Roenker DL Ball KK Clay OJ, Wadley VG. Cumulative meta-analysis of the relationship between useful field of view and driving performance in older adults: current and future implications. *Optom Vis Sci*, 82(2):724–731, 2005.
12. Lacherez P Hickson L Wood JM, Chaparro A. Useful field of view predicts driving in the presence of distracters. *Optom Vis Sci*, 89(4):373–381, 2012.
13. Roenker Daniel L. Miller Richard L. Griggs David S. Ball Karlene K., Beard Bettina L. Age and visual search: Expanding the useful field of view. *Journal of the Optical Society of America, A, Optics, Image and Science*, 5(12):2210–2219, 1988.
14. Owsley C Ball K. The useful field of view test: a new technique for evaluating age-related declines in visual function. *J Am Optom Assoc*, 64(1):71–79, 1993.
15. Cynthia Owsley, Karlene Ball, Michael E. Sloane, Daniel L. Roenker, and John R. Bruni. Visual/cognitive correlates of vehicle accidents in older drivers. *Psychology and aging*, 6(3):403–415, 1991.
16. Selander H. Dukic Willstrand T., Broberg T. Driving characteristics of older drivers and their relationship to the useful field of view test. *Gerontology*, 63(2):180–188, 2017.
17. Baldock M.R.J., Berndt A., and Mathias J.I. The functional correlates of older drivers’ on-road driving test errors. *Top Geriatr Rehabil*, 24:204–223, 2008.
18. BrainHQ. Ufov assessment manual. Technical report, 2011.
19. Belkacem Chikhaoui and Frank Gouineau. Towards automatic feature extraction for activity recognition from wearable sensors: A deep learning approach. In *2017 IEEE International Conference on Data Mining Workshops, ICDM Workshops 2017, New Orleans, LA, USA, November 18-21, 2017*, pages 693–702, 2017.
20. Belkacem Chikhaoui, Frank Gouineau, and Martin Sotir. A CNN based transfer learning model for automatic activity recognition from accelerometer sensors. In *Machine Learning and Data Mining in Pattern Recognition - 14th International Conference, MLDM 2018, New York, NY, USA, July 15-19, 2018, Proceedings, Part II*, pages 302–315, 2018.
21. Christopher Olah. Understanding lstm networks. Technical report, 2015.

22. Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.*, 15(1):1929–1958, 2014.
23. Martin Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, Manjunath Kudlur, Josh Levenberg, Rajat Monga, Sherry Moore, Derek G. Murray, Benoit Steiner, Paul Tucker, Vijay Vasudevan, Pete Warden, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. Tensorflow: A system for large-scale machine learning. In *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, pages 265–283, 2016.